Cubic Spline Hermite Interpolation with Linear Least Square Regression for Single Scanning Electron Microscope Image Signal-to-Noise Ratio Estimation

K. S. Sim, *Member, IAENG*, C. K. Toa, *Member, IAENG*, C. W. Ho

***Abstract*— The quality of the scanning electron microscope (SEM) images can be estimated using the signal-to-noise ratio (SNR). SNR is defined as the ratio of the desired signal to background noise. The noise which appeared in the SEM image is called Gaussian noise. Thus, if the SNR value is high, the image will have better quality since there is more useful information (the signal) than unwanted data (the noise). However, existing SNR estimation methods such as Nearest Neighbourhood (NN), Linear Interpolation (LI), and combination of Nearest Neighbourhood and Linear Interpolation unable to provide satisfactory results in estimating the SNR value. So, to prevent the loss of important information of SEM images, the novel SNR estimation method named Cubic Spline Hermite Interpolation with Linear Least Square Regression (CSHILLSR) has been proposed and formulated. The proposed method is compared with existing methods in terms of absolute error of SNR values, F-test, and Student's t-test. The result shows that the proposed method having a lower absolute error as compared to other methods and there is no significant difference between the actual and estimated SNR value at a 95% confidence level. This indicates that the proposed CSHILLSR able to provide better accuracy in estimation of SNR value as compare to the existing methods.**

***Index Terms*— Gaussian noise, SEM image, Absolute error, Signal-to-noise ratio (SNR), F-test, Student's t-test**

I. Introduction

SIGNAL to noise ratio (SNR) is a ratio that compares the desired signal to the background noise. It is used to determine the quality of the gray-scale images and the scanning electron microscope (SEM) images. If the image has good quality, the SNR value will high indicate that it has a higher level of the signal compared to the level of noise. At the same time, it also means that if the image has bad quality, it will have low SNR value due to the high level of noise to be presented in the image. The noise that typically existed in SEM images is derived from Gaussian white noise, which has uniformly distributed noise spectral density.

In this paper, the focus is on developing new SNR estimation methods to estimate the white Gaussian noise that existed in the grayscale SEM images. Due to the presence of white noise, SEM image information might be corrupted. As a result, this will cause some important information on SEM images to be lost. Hence, it is necessary to accurately measure the SNR of a single SEM image to provide useful information to the noise filter.

Three existing SNR estimation methods [1] named Nearest Neighbourhood (NN), Linear Interpolation (LI), and combination of Nearest Neighbourhood and Linear Interpolation are discussed. These three existing methods are then compared with the newly proposed method named Cubic Spline Hermite Interpolation with Linear Least Square Regression (CSHILLSR). In this paper, SEM images of various noise levels are applied to evaluate noise estimation accuracy.

This paper is arranged as follows. A few prevailing SNR estimation methods are discussed in the next section. The problem statement is elaborated in Section III. In Section IV, the detail of the proposed method is shown. In Section V, the results of the proposed method are presented. In Section VI, the discussion of the comparison between the proposed and existing methods are presented. In Section VII, a summary of the outcome of the proposed method is described.

II. Literature Review

Noise is being defined as an unwanted signal in the dictionary. Since it is an unwanted signal and is always present in the Scanning Electron Microscope (SEM) images, it will degrade image quality. So, the signal-to-noise ratio (SNR) parameter was invented and used to measure the level of noise in an image.

In the past, quite a few numbers of SNR estimation techniques had been proposed. The first technique was proposed by Frank & Al-Ali (1975) and is called the cross-correlation technique. This technique is applied according to the Fourier Transform theorem [2]. After the proposal had been made in 1975 by Frank & Al-Ali, the cross-correlation technique had been widely used to calculating the signal to noise ratio. For example, in 1982, Erasmus uses the cross-correlation technique to calculate the SNR in the digital image [3]. In 2000, Joy et al. applied a cross-correlation technique to estimate the SNR value of SEM image because of its critical dimension and its resolution [4]. But this

Manuscript received March 5, 2020; revised August 8, 2020. This work was supported in part by the Multimedia University Malaysia.

K. S. Sim, C. K. Toa, and C. W. Ho are with the Faculty of Engineering and Technology, Multimedia University, Jalan Ayer Keroh Lama, 75450 Melaka, Malaysia (email: toacheankhim@yahoo.com; kssim@mmu.edu.my)

technique required two images to be positioned perfectly with each other to predict the SNR values. The need for two images to calculate SNR value confines this technique to non-stored images, as a result, it limits its application. Besides, it also consumes a lot of time since two images need to position perfectly.

Due to the previous technique required two images to be positioned perfectly for estimating the SNR value, in 2004, Sim proposed a new technique named as autoregressive (AR) model [5]. The AR model uses only a single SEM image instead of two images to estimate the SNR value. It states that the output variable should depend linearly on the previous value. Besides, three methods under AR model named Linear Interpolation (LI) method, Nearest Neighbourhood (NN) method, and the combination of both methods are introduced. However, these three methods still need to improve the accuracy of SNR estimation.

Based on Fig. 1, LI method applies to the two nearest neighbour points before and after the noisy peak. These points are known as r(-1,y), r(-2,y), r(1,y), and r(2,y) as shown in Equation 1 and Equation 2.

$$r^{int}(0,y) = r(1,y) + [r(1,y)-r(2,y)] \tag{1}$$

$$r^{int}(0,y) = r(-1,y) + [r(-1,y)-r(-2,y)] \tag{2}$$

Equation 3 and Equation 4 are derived from Equation 1 and Equation 2.

$$r^{int}(0,y) = 2\,r(1,y) - r(2,y) \tag{3}$$

$$r^{int}(0,y) = 2\,r(-1,y) - r(-2,y) \tag{4}$$

where r(x,y) is the noise-free peak of the SEM image. Both equations will be used to predict the value of noise free peak. After that, Equation 5 is applied to obtain the SNR value.

$$SNR = \frac{r^{int}(0,y) - \mu^2}{r(0,y) - r^{int}(0,y)} \tag{5}$$

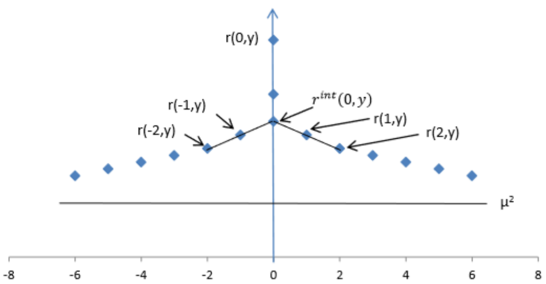

Fig. 1. Linear Interpolation method in estimating $r^{int}(0,y)$

The NN method is another method for estimating the SNR of an image. According to Fig. 2, the two closest points next to the noisy peak will be considered in this method. These two points are r(1,y) and r(-1,y) which can be assumed as the estimated noise-free peak $r^{NF}(0,y)$. The NN method is explained by Equation 6.

$$r^{NF}(0,y) = r(1,y) = r(-1,y) \tag{6}$$

Later, the value obtained from Equation 6 will be inserted into Equation 7 to find the SNR value.

$$SNR = \frac{r^{NF}(0,y) - \mu^2}{r(0,y) - r^{NF}(0,y)} \tag{7}$$

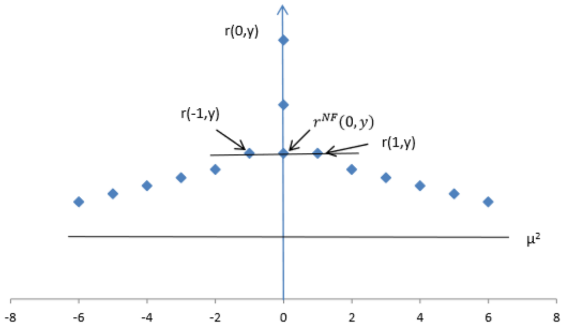

Fig. 2. Nearest Neighbourhood method in determining $r^{NF}(0,y)$

Next is the combination of LI method and NN method as shown in Equation 8. This method is developed by including only the good features found from these two methods.

$$r^{NN+int}(0,y) = \frac{r^{NN}(0,y) + r^{int}(0,y)}{2} \tag{8}$$

where $r^{NN+int}(0,y)$ is the estimated noise-free peak value. Equation 9 and Equation 10 are the basic forms of Equation 8 with the opposite number.

$$r^{NN+int}(0,y) = \frac{3r(1,y)+r(2,y)}{2} \tag{9}$$

$$r^{NN+int}(0,y) = \frac{3r(-1,y)+r(-2,y)}{2} \tag{10}$$

After that, Equation 9 and Equation 10 will be substituted into Equation 11 to calculate the SNR value.

$$SNR = \frac{r^{NN+int}(0,y) - \mu^2}{r(0,y) - r^{NN+int}(0,y)} \tag{11}$$

In 2005, Sim had formulated an algorithm to cascade the autoregressive model by using Lagrange Time Delay (LTD) [6]. This method is named as Mixed Lagrange Time Delay Estimation Autoregressive (MLTDEAR) model. In experiments involving different images, this method provides an optimum solution to the SNR estimation problem, while being affected by different noise environments. As a result, compared with the previous method included in the autoregressive model, this method proved to be more accurate and effective because it is not affected by addictive white noise. However, it is still affected by very low magnification and varying contrast of SEM images.

In 2008, Sim proposed another technique which is known as Shape-Preserving Cubic Hermite Autoregressive Moving

Average (SPCHAMA) [7]. This technique combines the Cubic Hermite curve, autoregressive model, and the average model to determine the SNR value. Besides, this technique is found to give optimum solutions in estimating SNR due to its robustness under different noise situations. Although both MLTDEAR and SPCHAMA are used to improve the AR model, these techniques bring a heavy computational burden.

In 2013, Gao and his partner had carried out some research about algorithms that depend on the linear regression by using the simulated images with numerous SNR values and real images [8]. Linear regression is a linear approach used to model the relationship that exists between the dependent variable and one or more independent variables. After the research, the result shows that linear regression is useful for noise estimation because it considers all data points in the graph. The definition of the linear regression method is shown in Equation 12.

$$y_i = Bx_i^2 + Gx_i + H + E \qquad (12)$$

Due to its characteristics in consideration of all data points in the graph, this method is used as a reference in our newly proposed method. Although a lot of research has been carried out, the existing methods have low accuracy in estimating the SNR value. Therefore, a new SNR estimation technique is introduced to improve the accuracy of SNR estimation.

III. PROBLEM STATEMENT

The image signal corrupted by noise is represented by Equation 13, where T(x,y) represents the actual signal, U(x,y) represents the White Gaussian noise and e(x,y) represents image function, respectively. This means that the image function is the sum of the actual signal and Gaussian white noise.

$$e(x,y) = T(x,y) + U(x,y) \qquad (13)$$

Based on Equation 13, a two-dimensional autocorrelation function (ACF) curve can be expressed in a one-dimensional ACF curve as shown in Fig. 3.

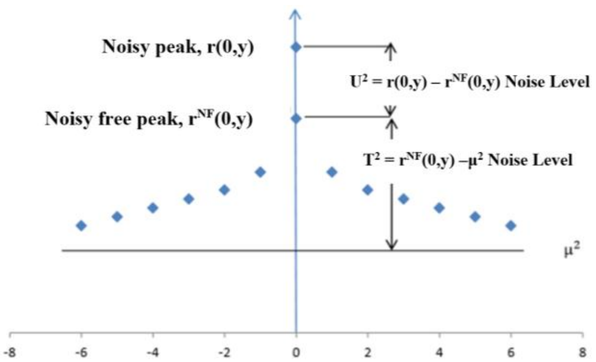

Fig. 3. The autocorrelation function (ACF) curve of SEM single image after added by White Gaussian Noise

Based on Figure 3, Equation 14 is used to determine the SNR value of the image.

$$SNR = \frac{r^{NF}(0,y) - \mu^2}{r(0,y) - r^{NF}(0,y)} \qquad (14)$$

where $r^{NF}(0,y)$ represents the noise-free peak, μ^2 represents the mean, and r(0,y) represents the noisy peak. The denominators are the differences between the noisy peak and the noise-free peak, while the numerators are the differences between the noise-free peak and mean. The mean value and the noisy peak value of the image can be obtained from the ACF curve.

To calculate the noise variance of the corrupted SEM image, Equation 15 can be used.

$$Noise\ Variance = \frac{r(0,y) - r^{NF}(0,y)}{Image\ Resolution} \qquad (15)$$

The denominator is represented by image resolution, where it is a pixel size of the SEM images. For instance, the image resolution can be from (256x256), (512x512), and so on. In this project, the image resolution (256x256) is used and fixed to maintain the consistency of the experiment [1].

If the SNR value and noise variance are known, the further work of reducing white noise will be simpler. Therefore, the main problem here is to accurately estimate the noise-free peak. Therefore, new methods have been developed to obtain the noise-free peak value more accurately. These methods will be discussed in the following chapter.

IV. METHODOLOGY

A new Cubic Spline Hermite Interpolation with Linear Least Square Regression (CSHILSR) is proposed to estimate the SNR value. The Cubic Spline Hermite Interpolation method makes the interpolating polynomial curve smoother and reduces the error. Moreover, to fit the model to the data, a linear least square regression method is applied. Equation 16 and Equation 17 shows the general equation of Cubic Spline.

$$P_i(x) = y_i + d_i\,(x-x_i) + e_i(x-x_i)^2 + f_i(x-x_i)^3 \qquad (16)$$

$$P_i(x) = y_{di} + d_i\,(x-x_i) + e_i(x-x_i)^2 + f_i(x-x_i)^3 \quad \text{for } x_i \le x \le x_{i+1} \qquad (17)$$

where i is the total data points, y_i represents the noise-free value at x_i, and $P_i(x)$ represents the cubic function.

By referring to Equation 16 and Equation 17, there are three coefficients which are d_i, e_i, and f_i needs to be identified. Thus, the boundary condition is applied to Equation 16 and the outcome is shown in Equation 18 and Equation 19.

$$P_i\,(x_i) = y_i \qquad \text{when } x=x_i \qquad (18)$$

$$P_i\,(x_{i+1}) = P_{i+1}\,(x_{i+1}) = y_{i+1} \qquad \text{when } x=x_{i+1} \qquad (19)$$

where P(x) is a cubic function with a smooth curve. Later, the P(x) formulation will be inserted into Equation 20 and Equation 21 for further calculation.

$$\frac{dP_i\,(x_{i+1})}{d(x_{i+1})} = \frac{dP_{i+1}\,(x_{i+1})}{d(x_{i+1})} \tag{20}$$

$$\frac{d^2P_i\,(x_{i+1})}{d(x^2{}_{i+1})} = \frac{d^2P_{i+1}\,(x_{i+1})}{d(x^2{}_{i+1})} \tag{21}$$

Equation 22 until Equation 24 have been developed based on Equation 20,

$$\frac{dP_i\,(x_{i+1})}{d(x_{i+1})} = d_i + 2e_ix_{i+1} - 2e_ix_i + 3f_ix_{i+1}{}^2 - 6f_ix_{i+1}x_i + 3f_ix_i{}^2 \tag{22}$$

$$\frac{dP_{i+1}\,(x_{i+1})}{d(x_{i+1})} = 0 \tag{23}$$

$$\frac{dP_i\,(x_{i+1})}{d(x_{i+1})} = \frac{dP_{i+1}\,(x_{i+1})}{d(x_{i+1})} \tag{24}$$

Thus, it yields and forms Equation 25.

$$d_i + 2e_ix_{i+1} - 2e_ix_i + 3f_ix_{i+1}{}^2 - 6f_ix_{i+1}x_i + 3f_ix_i{}^2 = 0 \tag{25}$$

Equation 26 until Equation 28 have been developed based on Equation 21,

$$\frac{d^2P_i\,(x_{i+1})}{d(x^2{}_{i+1})} = 2e_i + 6f_ix_{i+1} - 6f_ix_i \tag{26}$$

$$\frac{d^2P_{i+1}\,(x_{i+1})}{d(x^2{}_{i+1})} = 10 \tag{27}$$

$$\frac{d^2P_i\,(x_{i+1})}{d(x^2{}_{i+1})} = \frac{d^2P_{i+1}\,(x_{i+1})}{d(x^2{}_{i+1})} \tag{28}$$

Thus, it yields and forms Equation 29.

$$2e_i + 6f_ix_{i+1} - 6f_ix_i = 0 \tag{29}$$

In order to solve for f, g, and h, Equations 25, Equation 29 and Equation 30 are required to be used

$$P_i(x_{i+1}) = y_i + d_i\,(x_{i+1}-x_i) + e_i(x_{i+1}-x_i)^2 + f_i(x_{i+1}-x_i)^3 \tag{30}$$

Assume $x_i=x_1$ and $x_{i+1}= x_2$, the value of $P_1(x_1)$ can be obtained through Equation 17 which is equal to y_1. Besides, from Equation 18, since the y_2 value is known, the value of $P_2(x_2)$ and $P_1(x_2)$ can be obtained. After that, the $P_1(x_2)$ value will be substitute back into Equation 30 to get Equation 31.

$$P_1(x_2) = y_1 + d_1\,(x_2-x_1) + e_1(x_2-x_1)^2 + f_1(x_2-x_1)^3 \tag{31}$$

Since the x_1, x_2, $P_1(x_2)$, and y_1 are known and Equation 30 had been the yield, the unknown d, e, and f can be calculated by solving the Equations 25, Equation 29, and Equation 30 simultaneously.

For our proposed method, x_1 is assumed to be 128 while x_2 is assumed to be 129 since the SEM images with 256x256 pixels are used. After determining the d, e, and f, x=128.5 is substituted back into Equation 16 to estimate the $P_1(x)$ value which is the estimated noise-free value. This value is then substituted into Equation 32 to estimate the SNR value.

$$SNR = \frac{P_i(0,y) - \mu^2}{r(0,y) - P_i(0,y)} \tag{32}$$

where $P_i(0,y)$ represents noise-free value, r(0,y) represents the noisy peak and μ^2 represents mean value. Next, the Cubic Hermite curve is introduced to combine with the Cubic Spline curve as shown in Fig. 4.

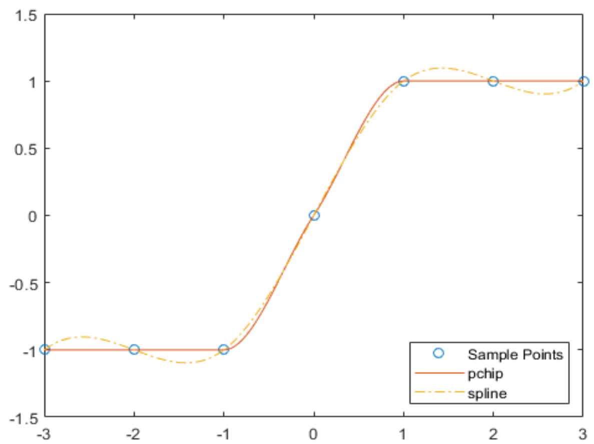

Fig. 4. The Cubic Spline and Cubic Hermite curve

Based on the graph, it shows that the Cubic Hermite curve has smaller oscillations as compared to the Cubic Spline curve. The reason for introducing the Cubic Hermite curve is to improve the accuracy since the Cubic Spline method keeps estimating the SNR value that is higher than the actual value. Besides, this method also mixed with the linear least square regression method as shown in Equation 33.

$$y_i = Bx_i^2 + Gx_i + H + E \tag{33}$$

where B, G, and H are constant-coefficient, x_i is the original SNR, E is the unexplained error, and y_i is the estimated SNR.

Since the linear least square regression method is mixed with the Cubic Spline interpolation method, Equation 34 until Equation 46 is developed to estimate the SNR value.

$$y_i = \frac{P_i(0,y) - \mu^2}{r(0,y) - P_i(0,y)} \tag{34}$$

$$\frac{P_i(0,y) - \mu^2}{r(0,y) - P_i(0,y)} = \text{Bxi2} + \text{Gxi} + \text{H} + \text{E} \tag{35}$$

$$\begin{bmatrix} y_1 \\ y_2 \\ \cdot \\ \cdot \\ y_N \end{bmatrix} = \begin{bmatrix} {x_1}^2 & x_1 & 1 \\ \vdots & \vdots & 1 \\ {x_N}^2 & x_N & 1 \end{bmatrix} \begin{bmatrix} B \\ G \\ H \end{bmatrix} + \begin{bmatrix} E_1 \\ E_2 \\ \cdot \\ \cdot \\ E_N \end{bmatrix} \tag{36}$$

$$let\ \varepsilon = y - XB \tag{37}$$

$$\frac{d}{d\beta}(\varepsilon^T \varepsilon) = \frac{d}{d\beta}(y - XB)^T (y - XB) = 0 \tag{38}$$

$$\frac{d}{d\beta}(y^T y - y^T XB + (XB)^T (XB) - (XB)^T y) = 0 \tag{39}$$

$$\frac{d}{d\beta}(y^T y - y^T XB + B^T X^T XB - y^T XB) = 0 \tag{40}$$

$$\frac{d}{d\beta}(y^T y - 2y^T XB + B^T X^T XB) = 0 \tag{41}$$

$$2X^T XB - 2X^T y = 0 \tag{42}$$

$$X^T XB = X^T y \tag{43}$$

$$\mathrm{B} = (\mathrm{X}^T\mathrm{X})^{-1}\mathrm{X}^T\mathrm{Y} \tag{44}$$

$$A = \begin{bmatrix} B \\ G \end{bmatrix} = (X^T X)^{-1} X^T Y \tag{45}$$

$$y_i = Bc_i^2 + Gx_i \tag{46}$$

where y_i is the estimated SNR, x_i is the original SNR, $P_i(0,y)$ is predicted noise-free, μ^2 is the mean of the image, and r(0,y) is the noisy peak. According to vector calculus, the derivation to minimize the unexplained error, E is expressed in Equation 47.

$$\left[\min_A \sum_{k=1}^{N} E_k^2 = \min_A E^T E\right] = \left[\frac{d}{dA}\sum_{k=1}^{N} E_k^2 = \frac{d}{dA} E^T E\right] \tag{47}$$

V. RESULTS

In this experiment, 1000 images had been used to perform the testing, but only six images of samples are selected and used here for the comparison purpose. These six images size was set to be in the pixel resolution of (256x256). A MATLAB program is used to perform the estimation of SNR value. The flowchart of Matlab is attached and shown in Fig. A1 in Appendix A.

To compare the performance of Cubic Spline Hermite Interpolation with Linear Least Square Regression (CSHILSR) with other existing methods namely Linear Interpolation (LI), Nearest Neighbourhood (NN), and the combination of LI and NN, three measurement formulations which are the absolute error of SNR, F-test, and Student's t-test have been carried out.

The SNR values of a single SEM image are calculated using Equation 48.

$$SNR = 10 \cdot log_{10}\left[\frac{\sum_0^{n_x-1}\sum_0^{n_y-1}[r(x,y)]^2}{\sum_0^{n_x-1}\sum_0^{n_y-1}[r(x,y) - t(x,y)]^2}\right] \tag{48}$$

where r(x,y) is the reference image and t(x,y) is the test image [9], [10]. The accuracy of the estimation methods is acquiring based on the absolute error between estimated and actual SNR value. A smaller absolute error value indicates higher accuracy of estimation method. The absolute error can be calculated using Equation 49.

Absolute Error=|Actual SNR- Estimated SNR| (49)

As for the F-test and Student's t-test, their formulation is shown in Equation 50 and Equation 51

$$t = \frac{\frac{(\sum d)}{n}}{\sqrt{\frac{\sum d^2 - \left(\frac{(\sum d)^2}{n}\right)}{(n-1)(n)}}} \tag{50}$$

where d represents the difference between the two groups and n represent the total number for a group.

$$f = \frac{S_{varaince\_1}}{S_{variance\_2}} \tag{51}$$

where $S_{varainc\_1}$ represent the largest population variance and $S_{varaince\_2}$ represent the smallest population variance. Fig. 5 to Fig. 10 illustrate six images used as a sample for this experiment.

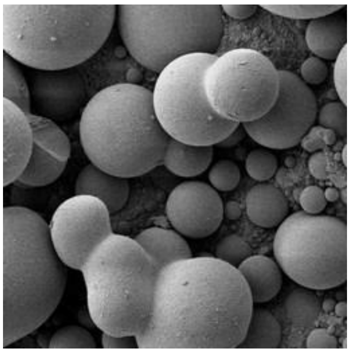
Fig. 5. Composite Material A

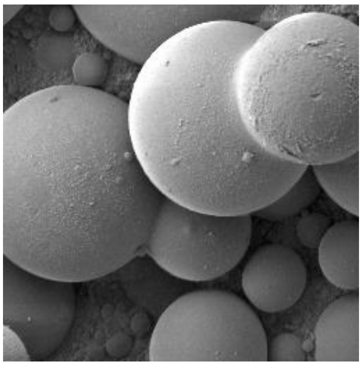
Fig. 6. Composite Material A with magnification

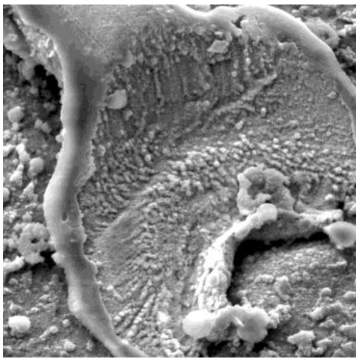

Fig. 7. Composite Material B

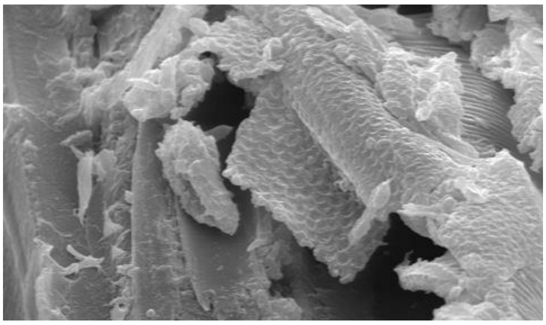

Fig. 8. Wood Fibre

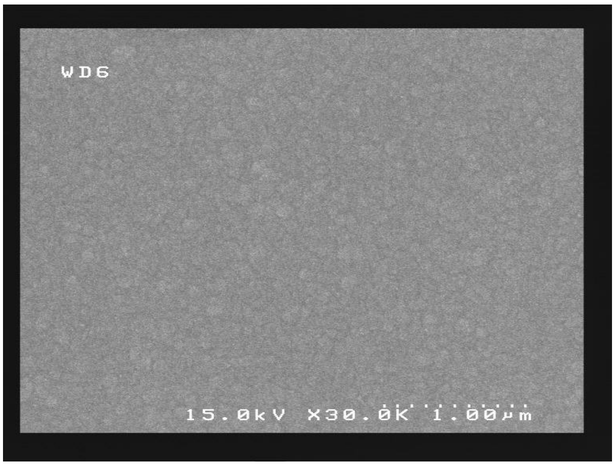

Fig. 9. Wafer with 30.00k speed

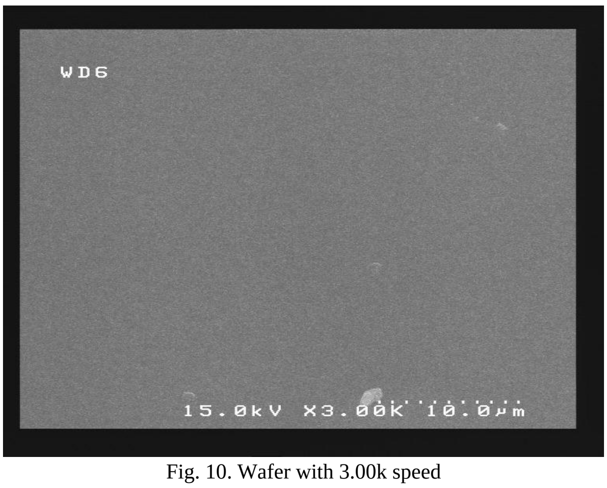

Fig. 10. Wafer with 3.00k speed

The peak of the autocorrelation function (ACF) curve of six SEM images is utilized to estimate the SNR value. The noise variance is ranged from 0.001 to 0.010 with 0.001 increments each time to perform SNR estimation. The white Gaussian Noise is the only noise applied to the SEM image. Besides, the graph of the absolute error between the estimated and actual SNR value is obtained using Equation 49.

The data of estimated and actual SNR is shown in Table AI, AIII, AV, AVII, AIX, and AXI in Appendix A. These experimental data correspond to the six SEM images from Figure 5 to Figure 10. The first column of the table refers to the noise level which ranged from 0.001 to 0.010. Meanwhile, the second column refers to the actual SNR value followed by the NN method, LI method, the combination of LI and NN methods, and the CSHILSR method. The unit used for SNR estimation is in decibel (dB).

As for the absolute error difference of various SNR estimation methods, the data are shown in Table AII, AIV, AVI, AVIII, AX, and AXII in Appendix A. The first column refers to the noise level, followed by absolute error values of NN method, LI method, the combination of NN method and LI method, and CSHILSR method. The absolute error differences results were plotted for clarity purpose as shown in Fig. 11 until Fig. 16. The explanation for each of the results is included.

Next, to test the similarity between the population variance of estimated and standard SNR value, an F-test formula with 95% confidence level is used. Furthermore, to determine whether there is a significant difference between the means of those two values, a Student’s t-test had been carried out. By performing the F-test and t-test using a data analysis tool in Excel, several parameters are calculated. In this test, there are two sets of data used in the F-test and t-test, which are the actual value and estimated value. For each set of data, the mean is calculated. Later, the F, F critical, t stat, t critical, and P value will be determined. The P value can be interpreted as the probability of getting a result equal to or higher than the actual situation if the null hypothesis is true. The P value is determined from the t-distribution table by using the t-statistic and degree of freedom [11]. The confidence level α, which is also called as significance level, is set at 0.05. For F-test, if F value is smaller than F critical one-tail value, the null hypothesis will not be rejected. While for t-test, two main conditions need to be fulfilled in order to avoid the null hypothesis being rejected. First, if the P value is more than α, the null hypothesis cannot be rejected. Secondly, if the t-stat is less than or equal to t-critical, the null hypothesis cannot be rejected.

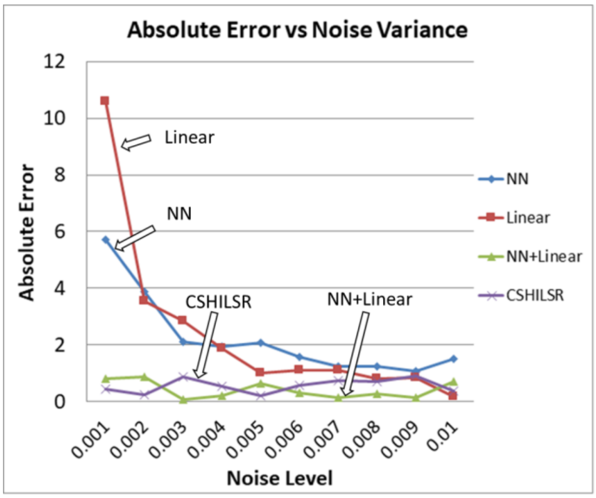

Fig. 11. Graph of absolute error vs noise variance of Fig. 5

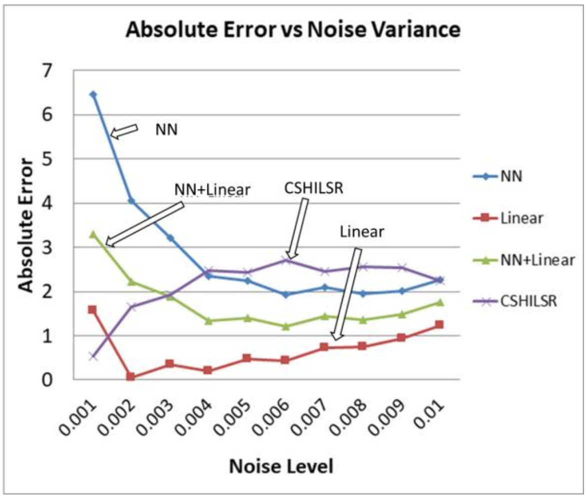

Fig. 12. Graph of absolute error vs noise variance of Fig. 6

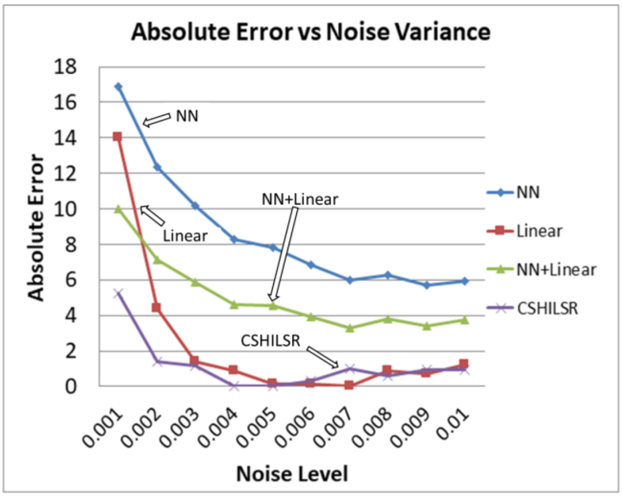

Fig. 13. Graph of absolute error vs noise variance of Fig. 7

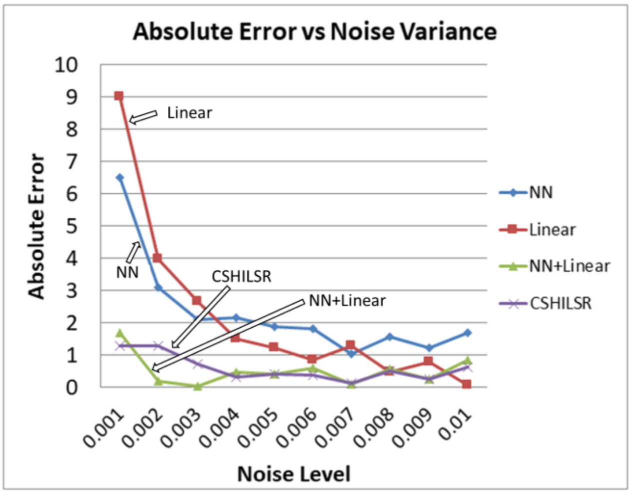

Fig. 14. Graph of absolute error vs noise variance of Fig. 8

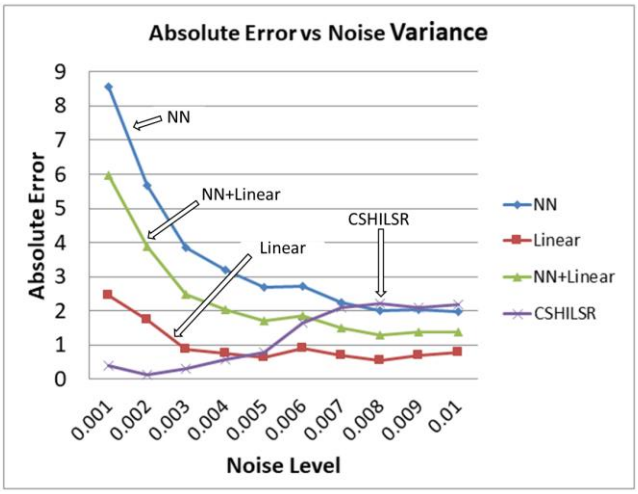

Fig. 15. Graph of absolute error vs noise variance of Fig. 9

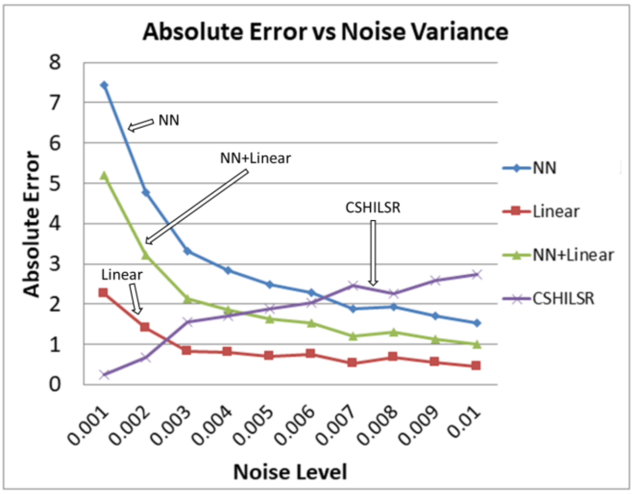

Fig. 16. Graph of absolute error vs noise variance of Fig. 10

The result of F test and t-test are shown in Table I, Table II, Table III, Table IV, Table V, and Table VI.

TABLE I
STUDENT'S F-TEST AND T-TEST OF CSHILSR FOR FIG.5

		Variable 1	*Variable 2*
	Mean	18.96039	19.52309
	Variance	38.79621	37.74246
F-test	F	1.02791955	
	P(F<=f) one-tail	0.48397663	
	F Critical one-tail	3.1788931	
t-test	t Stat	-0.20339	
	P(T<=t) one-tail	0.420555	
	t Critical one-tail	1.734064	
	P(T<=t) two-tail	0.841109	
	t Critical two-tail	2.100922	

TABLE II
STUDENT'S F-TEST AND T-TEST OF CSHILSR FOR FIG.6

		Variable 1	*Variable 2*
	Mean	20.59446	22.7477
	Variance	44.12883	36.73999
F-test	F	1.20111172	
	P(F<=f) one-tail	0.39466724	
	F Critical one-tail	3.1788931	
T-test	t Stat	-0.75718	
	P(T<=t) one-tail	0.229369	
	t Critical one-tail	1.734064	
	P(T<=t) two-tail	0.458737	
	t Critical two-tail	2.100922	

TABLE III
STUDENT'S F-TEST AND T-TEST OF CSHILSR FOR FIG.7

		Variable 1	*Variable 2*
	Mean	20.44144	20.05061
	Variance	45.83664	25.08631
F-test	F	1.82715762	
	P(F<=f) one-tail	0.19130743	
	F Critical one-tail	3.1788931	
T-test	t Stat	0.146756	
	P(T<=t) one-tail	0.442478	
	t Critical one-tail	1.734064	
	P(T<=t) two-tail	0.884956	
	t Critical two-tail	2.100922	

TABLE IV
STUDENT'S F-TEST AND T-TEST OF CSHILSR FOR FIG.8

		Variable 1	*Variable 2*
	Mean	16.75595	17.06848
	Variance	48.64075	40.80859
F-test	F	1.832656	
	P(F<=f) one-tail	0.39473	
	F Critical one-tail	3.314575	
T-test	t Stat	-0.1045	
	P(T<=t) one-tail	0.458965	
	t Critical one-tail	1.734064	
	P(T<=t) two-tail	0.917931	
	t Critical two-tail	2.100922	

TABLE V
STUDENT'S F-TEST AND T-TEST OF CSHILSR FOR FIG.9

		Variable 1	*Variable 2*
	Mean	18.38515	19.54648
	Variance	41.0307	30.13315
F-test	F	1.361646	
	P(F<=f) one-tail	0.326529	
	F Critical one-tail	3.178893	
T-test	t Stat	-0.43534	
	P(T<=t) one-tail	0.334246	
	t Critical one-tail	1.734064	
	P(T<=t) two-tail	0.668493	
	t Critical two-tail	2.100922	

TABLE VI
STUDENT'S F-TEST AND T-TEST OF CSHILSR FOR FIG.10

		Variable 1	*Variable 2*
	Mean	15.28942	17.10358
	Variance	41.54877	31.89591
F-test	F	1.302636	
	P(F<=f) one-tail	0.350045	
	F Critical one-tail	3.178893	
T-test	t Stat	-0.66941	
	P(T<=t) one-tail	0.255864	
	t Critical one-tail	1.734064	
	P(T<=t) two-tail	0.511727	
	t Critical two-tail	2.100922	

VI. Discussion

Based on Fig. 11 until Fig.16, it is noticeable that the CSHILSR method gave a lower value of absolute error as compared to other methods. For CSHILSR, the highest absolute error for all images is 5.2425. This value can be considered as the smallest when compared to the 16.8957 of NN method, 14.0260 of LI method and 10.0153 of the combination of LI method and NN method. The reason why the proposed method produces better results is that previous research done by Gao and his colleagues proved that linear least squares regression can improve SNR estimation [8]. Besides, the image composed of nonlinear features is also considered, so by combining the Cubic Spline curve and Cubic Hermite curve [7], they provide a platform that allows us to obtain more accurate SNR estimates.

As for Table I until Table VI, variable 1 is the actual SNR value, while Variable 2 is the estimated SNR value using CSHILSR. The null hypothesis of the F test and Student's t-test stated that there is no significant difference between the actual and estimated SNR value at 95% confidence interval. Based on the result of those tables, it shows that there is not much difference between the mean values of variable 1 and variable 2 since the percentage error is less than 12%. Furthermore, the F value is smaller than F critical one tail value, P(T<=t) two-tail value is larger than the α (0.05) and the t stat value is smaller than the t critical one-tail value. Therefore, the null hypothesis cannot be rejected.

VII. Conclusion

In a nutshell, Cubic Spline Hermite Interpolation with Linear Least Square Regression (CSHILSR) has been proven to have higher accuracy in estimating SNR compared to existing methods. Therefore, CSHILSR can become a useful method for noise filtering in the scanning electron microscope (SEM) images.

Appendix A

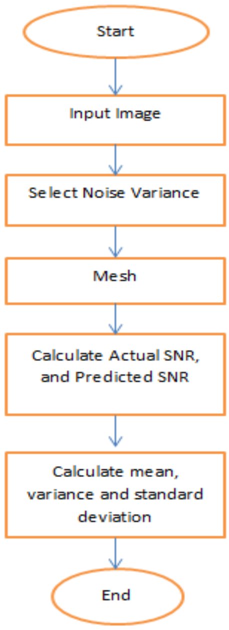

Fig. A1: Flowchart of Matlab program that used to calculate Actual SNR and Estimated SNR

TABLE AI
Estimated and Actual Signal-to-Noise ratio (SNR) of Fig. 5

Noise Level	Actual SNR (decibel)	Estimated SNR (decibel)			
		Nearest Neighbourhood (NN)	Linear interpolation (LI)	LI + NN	CSHILSRSNR
0.001	31.7316	26.0133	42.3098	30.9284	32.178
0.002	26.3205	22.4542	29.8508	25.4597	26.5567
0.003	22.02	19.8971	24.857	22.0772	22.8961
0.004	19.8828	17.9268	21.764	19.6741	20.4369
0.005	18.3667	16.2824	19.3645	17.7184	18.5845
0.006	16.589	15.0116	17.7054	16.282	17.1526
0.007	15.077	13.8366	16.168	14.9489	15.8176
0.008	14.0824	12.8379	14.8755	13.8172	14.7751
0.009	13.0014	11.9217	13.8462	12.8504	13.9173
0.010	12.5325	11.0283	12.6882	11.8348	12.9161

TABLE AII
Absolute error difference of various SNR estimation methods of Fig. 5

Noise Level	Estimated SNR (decibel)			
	Nearest Neighbourhood (NN)	Linear interpolation (LI)	LI + NN	CSHILSRSNR
0.001	5.7183	10.5782	0.8032	0.4464
0.002	3.8663	3.5303	0.8608	0.2362
0.003	2.1229	2.837	0.0572	0.8761
0.004	1.956	1.8812	0.2087	0.5541
0.005	2.0843	0.9978	0.6483	0.2178
0.006	1.5774	1.1164	0.307	0.5636
0.007	1.2404	1.091	0.1281	0.7406
0.008	1.2445	0.7931	0.2652	0.6927
0.009	1.0797	0.8448	0.151	0.9159
0.010	1.5042	0.1557	0.6977	0.3836

TABLE AIII
Estimated and Actual Signal-to-Noise ratio (SNR) of Fig. 6

Noise Level	Actual SNR (decibel)	Estimated SNR (decibel)			
		Nearest Neighbourhood (NN)	Linear interpolation (LI)	LI + NN	CSHILSRSNR
0.001	34.6433	28.1802	36.219	31.3441	35.1783
0.002	27.9524	23.8932	28.0058	25.7315	29.6089
0.003	24.2655	21.0477	23.9138	22.3795	26.1933
0.004	21.2467	18.8881	21.0436	19.9113	23.7206
0.005	19.4839	17.2323	19.0157	18.0884	21.909
0.006	17.7051	15.7784	17.2671	16.4992	20.4152
0.007	16.6383	14.543	15.9034	15.2045	19.0961
0.008	15.4979	13.5436	14.7561	14.1357	18.0497
0.009	14.5574	12.5453	13.6225	13.0733	17.1024
0.010	13.9541	11.6867	12.712	12.1902	16.2035

TABLE AIV
Absolute error difference of various SNR estimation methods of Fig. 6

Noise Level	Estimated SNR (decibel)			
	Nearest Neighbourhood (NN)	Linear interpolation (LI)	LI + NN	CSHILSRSNR
0.001	6.4631	1.5757	3.2992	0.535
0.002	4.0592	0.0534	2.2209	1.6565
0.003	3.2178	0.3517	1.886	1.9278
0.004	2.3586	0.2031	1.3354	2.4739
0.005	2.2516	0.4682	1.3955	2.4251
0.006	1.9267	0.438	1.2059	2.7101
0.007	2.0953	0.7349	1.4338	2.4578
0.008	1.9543	0.7418	1.3622	2.5518
0.009	2.0121	0.9349	1.4841	2.545
0.010	2.2674	1.2421	1.7639	2.2494

TABLE AV
Estimated and Actual Signal-to-Noise ratio (SNR) of Fig. 7

Noise Level	Actual SNR (decibel)	Estimated SNR (decibel)			
		Nearest Neighbourhood (NN)	Linear interpolation (LI)	LI + NN	CSHILSRSNR
0.001	34.4558	17.5601	48.4818	24.4405	29.2133
0.002	28.0757	15.6981	32.4431	20.9647	26.6981
0.003	24.3787	14.186	25.7806	18.5079	23.2013
0.004	21.2469	12.9419	22.1345	16.6482	21.2743
0.005	19.5927	11.7834	19.4536	15.0313	19.6161
0.006	17.6899	10.8281	17.5527	13.7628	18.0213
0.007	15.9649	9.95904	15.9995	12.6532	16.9776
0.008	15.321	9.04524	14.4593	11.5076	15.9109
0.009	14.0716	8.37063	13.3716	10.674	15.033
0.010	13.6172	7.6904	12.3953	9.87877	14.5602

TABLE AVI
Absolute error difference of various SNR estimation methods of Fig. 7

Noise Level	Estimated SNR (decibel)			
	Nearest Neighbourhood (NN)	Linear interpolation (LI)	LI + NN	CSHILSRSNR
0.001	16.8957	14.0260	10.0153	5.2425
0.002	12.3776	4.3674	7.1110	1.3776
0.003	10.1927	1.4019	5.8708	1.1774
0.004	8.3050	0.8876	4.5987	0.0274
0.005	7.8093	0.1391	4.5614	0.0234
0.006	6.8618	0.1372	3.9271	0.3314
0.007	6.0059	0.0346	3.3117	1.0127
0.008	6.2758	0.8617	3.8134	0.5899
0.009	5.7010	0.7000	3.3976	0.9614
0.010	5.9268	1.2219	3.7384	0.9430

TABLE AVII
Estimated and Actual Signal-to-Noise ratio (SNR) of Fig. 8

Noise Level	Actual SNR (decibel)	Estimated SNR (decibel)			
		Nearest Neighbourhood (NN)	Linear interpolation (LI)	LI + NN	CSHILSRSNR
0.001	30.5335	24.0231	39.5206	28.853	31.8294
0.002	23.3472	20.2498	27.322	23.1704	24.6165
0.003	19.7342	17.6383	22.3979	19.7545	20.4431
0.004	17.7983	15.6446	19.2993	17.3262	18.1014
0.005	15.919	14.0373	17.1549	15.4966	16.3435
0.006	14.5365	12.7109	15.3721	13.9336	14.9101
0.007	12.5516	11.5265	13.8509	12.6404	12.6696
0.008	11.9996	10.4225	12.485	11.4185	11.4939
0.009	10.7724	9.55402	11.5445	10.5184	10.5321
0.010	10.3672	8.68846	10.4173	9.53138	9.7452

TABLE AVIII
Absolute error difference of various SNR estimation methods of Fig. 8

Noise Level	Estimated SNR (decibel)			
	Nearest Neighbourhood (NN)	Linear interpolation (LI)	LI + NN	CSHILSRSNR
0.001	6.5104	8.9871	1.6805	1.2959
0.002	3.0974	3.9748	0.1768	1.2693
0.003	2.0959	2.6637	0.0203	0.7089
0.004	2.1537	1.5010	0.4721	0.3031
0.005	1.8817	1.2359	0.4224	0.4245
0.006	1.8256	0.8356	0.6029	0.3736
0.007	1.0251	1.2993	0.0888	0.1180
0.008	1.5771	0.4854	0.5811	0.5057
0.009	1.2184	0.7721	0.2540	0.2403
0.010	1.6787	0.0501	0.8358	0.6220

TABLE AIX
Estimated and Actual Signal-to-Noise ratio (SNR) of Fig. 9

Noise Level	Actual SNR (decibel)	Estimated SNR (decibel)			
		Nearest Neighbourhood (NN)	Linear interpolation (LI)	LI + NN	CSHILSRSNR
0.001	31.6159	23.0622	29.1661	25.6382	31.2338
0.002	25.7421	20.0593	24.0124	21.8456	25.8803
0.003	21.7464	17.8884	20.8684	19.2759	22.0428
0.004	19.3459	16.1623	18.5993	17.3159	19.9223
0.005	17.3652	14.669	16.7433	15.6618	18.1372
0.006	16.1121	13.3993	15.2076	14.2717	17.7513
0.007	14.549	12.3021	13.8508	13.0545	16.6403
0.008	13.3584	11.3589	12.8273	12.0744	15.5584
0.009	12.4306	10.3988	11.7377	11.0537	14.5327
0.010	11.5859	9.61057	10.8007	10.1949	13.7657

TABLE AX
Absolute error difference of various SNR estimation methods of Fig. 9

Noise Level	Estimated SNR (decibel)			
	Nearest Neighbourhood (NN)	Linear interpolation (LI)	LI + NN	CSHILSRSNR
0.001	8.5537	2.4498	5.9777	0.3821
0.002	5.6828	1.7297	3.8965	0.1382
0.003	3.8580	0.8780	2.4705	0.2964
0.004	3.1836	0.7466	2.0300	0.5764
0.005	2.6962	0.6219	1.7034	0.7720
0.006	2.7128	0.9045	1.8404	1.6392
0.007	2.2469	0.6982	1.4945	2.0913
0.008	1.9995	0.5311	1.2840	2.2000
0.009	2.0318	0.6929	1.3769	2.1021
0.010	1.9753	0.7852	1.3910	2.1798

TABLE AXI
Estimated and Actual Signal-to-Noise ratio (SNR) of Fig. 10

Noise Level	Actual SNR (decibel)	Estimated SNR (decibel)			
		Nearest Neighbourhood (NN)	Linear interpolation (LI)	LI + NN	CSHILSRSNR
0.001	28.5549	21.1098	26.2908	23.3663	28.8101
0.002	22.58	17.8131	21.1772	19.3643	23.2448
0.003	18.718	15.403	17.8787	16.5755	20.2864
0.004	16.3812	13.5514	15.5695	14.5204	18.0977
0.005	14.496	11.9982	13.7824	12.8615	16.3789
0.006	12.9243	10.6294	12.1722	11.3811	14.9459
0.007	11.3582	9.48764	10.829	10.1447	13.8166
0.008	10.4133	8.48713	9.74075	9.10306	12.6683
0.009	9.21842	7.52281	8.67353	8.08989	11.7947
0.010	8.2499	6.70682	7.78756	7.24056	10.9924

TABLE AXII
Absolute error difference of various SNR estimation methods of Fig. 10

Noise Level	Estimated SNR (decibel)			
	Nearest Neighbourhood (NN)	Linear interpolation (LI)	LI + NN	CSHILSRSNR
0.001	7.4451	2.2641	5.1886	0.2552
0.002	4.7669	1.4028	3.2157	0.6648
0.003	3.3150	0.8393	2.1425	1.5684
0.004	2.8298	0.8117	1.8608	1.7165
0.005	2.4978	0.7136	1.6345	1.8829
0.006	2.2949	0.7521	1.5432	2.0216
0.007	1.8706	0.5292	1.2135	2.4584
0.008	1.9262	0.6725	1.3102	2.2550
0.009	1.6956	0.5449	1.1285	2.5763
0.010	1.5431	0.4623	1.0093	2.7425

REFERENCES

[1] K. S. Sim, M. E. Nia, C. P. Tso, and W. K. Lim, “Performance of new signal to-noise ratio estimation for SEM images based on single image noise cross correlation,” Journal of Microscopy, vol. 248, no. 2, pp 120–128, 2012.
[2] J. Frank and L. Al-Ali, “Signal-to-noise ratio of electron micrographs obtained by cross correlation,” Nature, vol. 256, no. 5516, pp 376–379, 1975.
[3] S. J. Erasmus, “Reduction of noise in TV rate electron microscope images by digital filtering,” Journal of Microscopy, vol. 127, no. 1, pp 29–37, Jul. 1982.
[4] D. C. Joy, Y. Ko, and J. J. Hwu, “Metrics of resolution and performance for CD-SEMs,” Proc. SPIE 3998, Metrology, Inspection, and Process Control for Microlithography XIV, pp 108–114, 2000.
[5] K. S. Sim and N. S. Kamel, “Image signal-to-noise ratio estimation using the autoregressive model,” Scanning, vol. 26, no. 3, pp 135–139, 2004.
[6] K. S. Sim, H. T. Chuah, and C. Zheng, “Performance of a mixed Lagrange time delay estimation autoregressive (MLTDEAR) model for single-image signal-to-noise ratio estimation in scanning electron microscopy,” Journal of Microscopy, vol. 219, no. 1, pp 1–17, 2005.
[7] K. S. Sim, M. Y. Wee, and, W. K. Lim, “Image signal-to-noise ratio estimation using shape-preserving piecewise cubic Hermite autoregressive moving average model,” Microscopy Research and Technique, vol. 71, pp 710-720, 2008.
[8] L. R. Gao, Q. Du, B. Zhang, W. Yang, and Y. F. Wu, “A Comparative Study on Linear Regression-Based Noise Estimation for Hyperspectral Imagery,” IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 6, no. 2, pp 488–498, 2013.
[9] K. S. Sim and C. K. Toa, “Nonlinear Spatial Domain First Order Moment Estimation in Magnitude Resonance Imaging Data,” International Conference on Robotics, Automation and Sciences, pp 1-5, 2016.
[10] C. K. Toa, K. S. Sim, Z. Y. Lim, and C. P. Lim, “Magnetic Resonance Imaging Noise Filtering using Adaptive Polynomial-Fit Non-Local Means,” Engineering Letters, vol. 27, no. 3, pp 527–540, 2019.
[11] G. D. Ruxton, “The unequal variance t-test is an underused alternative to Student's t-test and the Mann–Whitney U test. Behavioral Ecology,” Behavioral Ecology, vol. 17, no. 4, pp 688-690, 2006.